Audio-visual interaction and perceptual assessment of water features used over road traffic noise

Citation for published version:
Galbrun, L & Calarco, FMA 2014, 'Audio-visual interaction and perceptual assessment of water features used over road traffic noise', *Journal of the Acoustical Society of America*, vol. 136, no. 5, pp. 2609-2620.
https://doi.org/10.1121/1.4897313

Digital Object Identifier (DOI):
10.1121/1.4897313

Published In:
Journal of the Acoustical Society of America

Audio-visual interaction and perceptual assessment of water features used over road traffic noise[a)]

Laurent Galbrun[b)] and Francesca M. A. Calarco
School of the Built Environment, Heriot-Watt University, Edinburgh EH14 4AS, United Kingdom

(Received 17 April 2014; revised 15 August 2014; accepted 22 September 2014)

This paper examines the audio-visual interaction and perception of water features used over road traffic noise, including their semantic aural properties, as well as their categorization and evocation properties. The research focused on a wide range of small to medium sized water features that can be used in gardens and parks to promote peacefulness and relaxation. Paired comparisons highlighted the inter-dependence between uni-modal (audio-only or visual-only) and bi-modal (audio-visual) perception, indicating that equal attention should be given to the design of both stimuli. In general, natural looking features tended to increase preference scores (compared to audio-only paired comparison scores), while manmade looking features decreased them. Semantic descriptors showed significant correlations with preferences and were found to be more reliable design criteria than physical parameters. A principal component analysis identified three components within the nine semantic attributes tested: "emotional assessment," "sound quality," and "envelopment and temporal variation." The first two showed significant correlations with audio-only preferences, "emotional assessment" being the most important predictor of preferences, and its attributes naturalness, relaxation, and freshness also being significantly correlated with preferences. Categorization results indicated that natural stream sounds are easily identifiable (unlike waterfalls and fountains), while evocation results showed no unique relationship with preferences.
 [http://dx.doi.org/10.1121/1.4897313]

PACS number(s): 43.50.Rq, 43.50.Qp, 43.50.Fe [KVH] Pages: 2609–2620

I. INTRODUCTION

The use of water generated sounds is increasingly being considered in the built environment as a complement to conventional noise mitigation strategies, due to the inherent positive and relaxing qualities of water sounds,[1] and their ability to mask noise.[2–9] This concept has obtained greater interest over the last decade, as it has been shown that pleasant sounds (e.g., water, bird songs, bells, and wind in trees) can play an important role in acoustic comfort.[10] In particular, water sounds have often been identified as the best sounds to use for enhancing the urban soundscape in view of reducing stress and improving quality of life.[1,3]

The understanding of how to design water features from an audio-visual perspective is however still limited. Different designs can greatly affect the way in which water features are perceived both aurally and visually, but only a few recent studies have examined the physical and perceptual properties of water features in view of providing evidence-based design solutions.[2,6–9] The approaches and methodologies used so far have largely focused on acoustical preferences, with limited consideration given to visual preferences. The results presented in this paper aim to fill this gap by investigating how the acoustical and visual design of water features can affect preferences and perception of the water sounds, when used over road traffic noise to promote peacefulness and relaxation. Furthermore, the research also examines the subjective categorization and evocation properties of the water sounds. This research follows from previous work by Galbrun and Ali.[7] In particular, the study examines water features and streams of small to medium size (waterfalls, fountains, jets, a cascade, and a natural stream) which can typically be found in gardens and parks.

The soundscape approach (physical characteristics and mental perception of the aural environment[11]) has been extensively used to analyze water features, but only a few recent studies have examined the physical and perceptual properties of water sounds in detail.[2–9] These studies have focused on the use of water sounds over road traffic noise, the latter being the major urban noise annoyance. Most of this recent research concentrated on acoustical preferences of water sounds in the context of tranquility and relaxation,[2,3,6,7,9] and the main findings can be summarized as follows: (1) Water sounds are effective maskers of road traffic noise at mid-high frequencies but not at low frequencies,[2,7] although improvements in tranquility can be obtained even for low levels of masking;[2] (2) Stream sounds tend to be preferred to fountain sounds,[7] which are in turn preferred to waterfall sounds;[7,8] (3) Water sounds which are perceived to be manmade tend not to be liked;[2] (4) Water tends to be the preferred impact material, while flat surfaces made of hard materials are poorly rated;[7] (5) The preferred level of water sounds is similar or not less than 3 dB below the road traffic noise level;[3,7,12] (6) Water sounds with low sharpness tend to promote calmness and relaxation;[6,7] (7) Preference scores tend to increase when visual images are included in the

[a)]Portions of this work were presented in "Audio-visual interaction and perception of waterscapes used in outdoor environments," Proceedings of Inter-noise 2013, Innsbruck, Austria, September 2013.

[b)]Author to whom correspondence should be addressed. Electronic mail: l.g.u.galbrun@hw.ac.uk

tests;[2,6] (8) The visual stimulus is particularly important at low levels of road traffic noise, while the acoustic stimulus dominates perception at high levels of road traffic noise.[9] (9) Still water is visually calming, but higher visual preferences tend to occur for upward jets or a mix of different kinds of water features.[13] These findings indicate that recent research has mainly advanced the understanding of acoustical preferences of water features, while there is limited knowledge about the audio-visual interaction of water features.

Studies have shown that the degree of matching (congruence or coherence) between auditory and visual information is a significant component in sound ratings,[14,15] but research has been limited to the impact of features located within different environments,[6,16] rather than the impact of the water displays on their own. This is a limitation in relation to the work presented here, as the setting in which water features are placed can greatly influence preferences: For example, greenery is preferred to buildings, as the percentage of natural features at a location is a key factor influencing tranquility[17] and preferences.[9] Hong and Jeon[9] also found that the visual effects of vegetation on esthetic preference were significant, while those of water features were relatively small.

In the research presented here, the visual impact of water features' displays is examined by using images in which the displays are placed over the same natural background. This allows analyzing the effect of the water displays rather than their background. Additionally, the qualitative characterization of water sounds is also considered through semantic analysis and targeted questions, to gain a better understanding of the factors affecting water sounds' preferences. The findings obtained allowed identifying which water sounds and visual displays (of small to medium sized water features) are more suitable for improving peacefulness and relaxation within gardens and parks where road traffic noise is audible.

II. METHODS

The water features considered here are based on previous work by Galbrun and Ali[7] in which a variety of waterfalls, fountains, cascades, and jets were tested in the laboratory under controlled conditions, and one shallow stream was tested in the field. For the current study, 10 out of 12 water sounds have been selected from this pool of data (Table I). The selected features are representative of a wide range of acoustical and visual conditions and can each be classified into one of the following three categories: waterfall, fountain [with upward jet(s)], and stream. It can be noted that LJT is listed in Table I both as objective category 2 (fountain) and category 3 (stream). LJT is in fact an upward jet (i.e., a fountain), but it has been defined in previous research[7] as a stream type of sound because of its very shallow and irregular distribution of water (due to the low pressure present at its large nozzle and unsteady operation of the pump). Compared to the 12 features examined by Galbrun and Ali,[7] the selection excludes hard impact surfaces for the 37 jets fountain and the waterfall with small holes, as these were poorly rated when compared with water as the impact material.[7] Table I lists properties of the water features, including acoustic and psychoacoustic parameters of the sounds normalized to 55 dBA, both for water sounds and for road traffic noise. In line with previous work,[7] dense road traffic with low temporal variability has been used, as it is representative of a real case scenario where masking by small to medium sized water features could be used (e.g., in a garden or park). The road traffic noise was measured and recorded in a field at 200 m from the center of a busy motorway (M8 Edinburgh-Glasgow, UK; same as Ref. 7).

Acoustic parameters given in Table I were measured using an integrating-averaging sound level meter Brüel & Kjaer type 2250 (Naerum, Denmark), with a data averaging period of 20 s. Audio recordings of 20 s were carried out with a digital sound recorded Zoom H4n connected to Brüel & Kjaer type 4190 1/2 in. microphones, which were in turn attached outside the ears of a lightweight dummy head Sennheiser MKE 2002. The recordings were input into the MATLAB Software PsySound3 to compute the sharpness,[18] roughness,[19] and pitch strength[20] given in Table I. These psychoacoustic parameters were calculated from the 7 s normalized audio files used in the audio-only and audio-visual tests (see Sec. II A). Further details about the measurement procedures and the laboratory rig structure used for tests can be found in Ref. 7.

TABLE I. Properties of water sounds and road traffic noise used in the tests, including acoustic and psychoacoustic parameters of the sounds normalized to 55 dBA. Category numbers: 1 = waterfall, 2 = fountain, 3 = stream. The numbers in italic were calculated from sounds including both road traffic noise and water sounds. Fountain extensions and jets were placed at water level; the large jet had a nozzle's diameter of 25 mm, and the narrow jet had a nozzle's diameter of 10 mm.

Sound code	Water feature type & Category number	Impact material	Flow rate (l/min)	Height (m) & Width (m)	L_{A10}-L_{A90} (dB)		L_{Ceq}-L_{Aeq} (dB)		Sharpness (acum)		Roughness (asper)		Pitch strength	
PEW	Plain edge waterfall – 1	Water	120	1.0 – 1.0	1.1	*1.4*	−0.3	*2.8*	1.98	*1.70*	0.03	*0.04*	0.04	*0.07*
SEW	Sawtooth edge waterfall – 1	Water	30	0.5 – 1.0	1.0	*1.6*	−0.1	*2.7*	1.92	*1.59*	0.05	*0.05*	0.10	*0.07*
SHW	Small holes waterfall – 1	Water	30	0.5 – 1.0	0.7	*1.4*	−1.0	*2.5*	2.23	*1.71*	0.02	*0.04*	0.09	*0.08*
FTW	Fountain (37 jets) – 2	Water	30	-	1.4	*1.5*	−0.9	*2.7*	2.21	*1.67*	0.07	*0.08*	0.10	*0.08*
DF	Dome fountain – 2	Water	40	-	1.2	*1.4*	−1.0	*2.5*	2.16	*1.70*	0.05	*0.05*	0.11	*0.08*
FF	Foam fountain – 2	Stones & boulders	30	-	2.3	*1.6*	−0.2	*2.8*	1.91	*1.61*	0.09	*0.09*	0.05	*0.07*
LJT	Large jet – 2/3	Water	15	-	4.9	*2.1*	4.9	*2.9*	1.73	*1.42*	0.28	*0.19*	0.08	*0.07*
NJT	Narrow jet – 2	Water	15	-	1.9	*1.6*	−0.9	*2.5*	2.09	*1.67*	0.19	*0.16*	0.07	*0.08*
CA	Cascade (4 steps) – 3	Stones (pebbles)	15	-	1.2	*1.4*	−1.3	*2.7*	2.21	*1.71*	0.10	*0.09*	0.05	*0.08*
ST	Stream – 3	Stones and water	N/A	-	2.4	*1.7*	1.4	*2.5*	1.99	*1.61*	0.29	*0.21*	0.06	*0.08*
RTN	Road Traffic Noise	-	-	-	2.7		7.8		1.04		0.03		0.09	

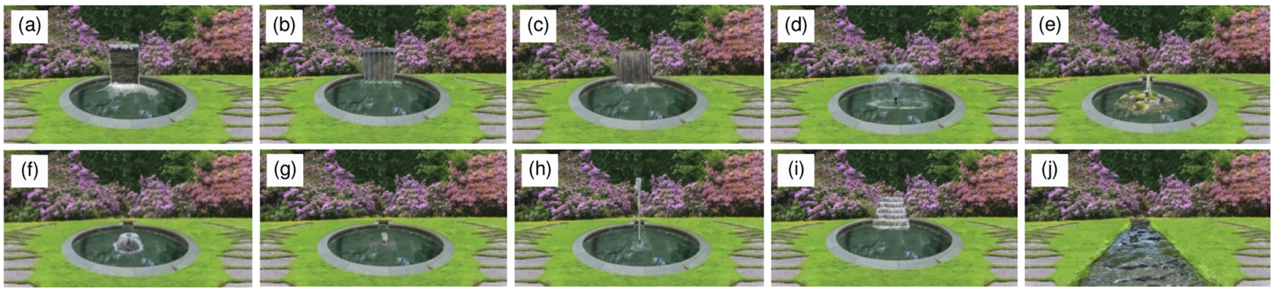

FIG. 1. (Color online) Visual materials used in the experiments. (a) PEW, waterfall with a plain edge, (b) SEW, waterfall with a sawtooth edge, (c) SHW, waterfall with an edge made of small holes, (d) FTW, fountain with 37 upward jets, (e) FF, foam fountain, (f) DF, dome fountain, (g) LJT, large jet, (h) NJT, narrow jet, (i) CA, cascade, and (j) ST, natural stream.

Visual material consisted of photomontages (Fig. 1) in which different water features' displays were placed over the same natural background using photographs. These images were developed using the photo editing software Adobe Photoshop CS3. In Fig. 1, the background is a Heriot-Watt University garden, and the water displays are almost identical to those tested in the laboratory (edited version of the photographs taken in the laboratory), with the exception of the stream tested in the field.[7]

A. Paired comparisons

Three different paired comparison tests were carried out to assess the interaction between the acoustical and visual stimuli: An audio-only test, a visual-only test, and an audio-visual test. The paired comparisons produced ordinal data that was appropriate for ranking preferences. This method has often been used in soundscape research[3,5,7,12] and was preferred to rating scales because of its simplicity and greater accuracy:[21] paired comparisons guaranteed a definite and more accurate ranking order through forced choice, unlike rating scales that would have allowed subjects to give identical scores to different waterscapes.

In all the tests, subjects had to imagine that they were relaxing in a balcony or garden where they could hear road traffic noise from a nearby motorway as well as a water feature (same as Refs. 2 and 7). In the tests, binaural signals were played back from a computer through closed headphones (Beyerdynamic DT 150), and the images of each water feature were presented to subjects on a large 27 in. light emitting diode (LED) monitor (Samsung LS27A350), as shown in Fig. 2. The tests were carried out in the anechoic chamber of Heriot-Watt University, a highly insulated space with a background noise level of 21 dBA during tests (including noise from the computer used). All the water sounds and road traffic noise were played at the same sound pressure level, as a difference of 0 dB between water sounds and traffic noise tends to be preferred.[6,7,12] The level used was 55 dBA, as it characterizes an outdoor environment that can significantly benefit from the use of water features, being not too quiet (no need for masking) and not too noisy (masking irrelevant for relaxation). Furthermore, the use of either 55, 70, or 75 dBA has been shown to have little effect on the preference findings of water sounds played over road traffic noise;[9,12] the higher levels just limited the importance of the visual stimulus.[9]

For each paired comparison of the audio-only test (no images used), subjects had to select the sound which they found more peaceful and relaxing (i.e., more tranquil[2]). In the paired comparisons of the visual-only test (no sounds played), subjects had to select the image (Fig. 1) that they preferred to look at. In the audio-visual test, subjects could hear and see pairs of water features, and they had to select the feature which they preferred in terms of peacefulness and relaxation. Each of the three tests included 45 paired comparisons, with 10 additional repetitions in the audio-only test, in view of identifying subjects' consistency. Each paired comparison consisted of 7 s of item 1 (sound or image or both), 1 s of silence/blank slide, 7 s of item 2, and 3 s of silence/blank slide, before the next pair was played. For statistical validity, the sequences of paired comparisons were randomized (different for each subject), and inconsistent subjects were removed from the analysis of results. For each test, five paired comparisons were initially played for familiarization with the methods. Once the subject was clear about the procedure, the actual test could begin. Each test consisted of ten paired comparisons played in an automated sequence, after which the subject was free to take a break before continuing, in order to maintain a high concentration level. The paired comparisons of the audio-only test lasted around 20 min and were followed by further qualitative analysis of the water sounds that lasted 20–30 min (see Sec. II B). The visual-only and audio-visual paired comparison tests each lasted around 20 min per subject and were not run straight after the audio-only test, to avoid fatigue. A further question was also asked for each image used in the tests (see Sec. II B), adding 5–10 min more to the total duration of the tests.

All the paired comparison data was normalized: Scale values were obtained by normalizing the number of times a waterscape was preferred [0 to 9 (number of paired comparisons per waterscape)] to an arbitrary -2 to $+2$ scale (as previously done

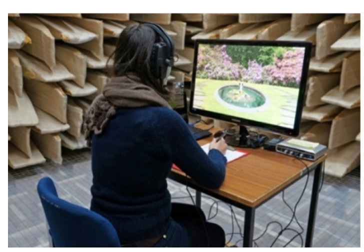

FIG. 2. (Color online) Laboratory setting used for the tests.

in Ref. 7). This provided a simple scale where positive values indicated waterscapes preferred a majority of times (+2 meaning always preferred), and negative values indicated waterscapes preferred a minority of times (−2 meaning never preferred). This normalization was carried out for each subject's data, and arithmetic averages were calculated across all subjects to find the values shown in Sec. III A.

B. Semantic assessment, categorization, and evocation

Water sounds' qualitative properties were also examined through the use of semantic scales and targeted questions aimed at categorization and evocation. After the audio-only paired comparisons, the ten water sounds with road traffic noise were played individually. Subjects could listen to each sound as many times as they wanted, and for each of these, they had to answer a questionnaire. This consisted mainly of semantic differential questions based on a five-point verbal scale, with nine attributes and antonymous adjectives that were carefully selected from a review of previous work.[22–25] These attributes/adjectives were relaxation (relaxing-stressful), naturalness (natural-artificial), freshness (refreshing-weary), familiarity (familiar-unfamiliar), perceived sharpness (sharp-flat), perceived roughness (rough-smooth), speed (fast-slow), envelopment (enveloping-directional), and temporal variation (unsteady-steady). These were selected to represent both emotional/qualitative attributes, as well as physical attributes that could be correlated with acoustical and psychoacoustical parameters. In addition to the hypothesized importance of relaxation and naturalness, freshness was selected because of previous work[6] that identified it as important in water sounds' preferences (a refreshing water sound being intended as energetic and cool). Familiarity was also included to examine prior experience and evocation, while the remaining components were selected to study correlations with acoustic and psychoacoustic parameters [$L_{A10} - L_{A90}$ (i.e., temporal variation), sharpness, roughness], as well as the perception of acoustic envelopment and water flow rate (speed). Descriptions and examples of the psychoacoustical and acoustical parameters were given to the subjects to ensure reliable collection of data. Regarding five-point scale questions, subjects were asked, for example, "*How relaxing is this sound?*" and had to tick one of the following answers: *very relaxing*, *relaxing*, *neither relaxing nor stressful*, *stressful*, *very stressful*.

After having completed all the semantic questions, subjects had to answer which type of water feature the sound made them think of (categorization): *waterfall*, *fountain*, *natural stream* or *none of these*. Evocation was also examined by asking the following questions: "*If the sound evokes anything to you, please explain what it makes you think of*" (open answer), "*Does this sound make you think of a manmade water feature? (e.g., water falling into a drain/container or a tap)*" (*yes* or *no*) and "*Does this sound make you think about rainfall?*" (*yes* or *no*).

All of the above tests concentrated on the acoustical perception of the water sounds. In addition, one visual assessment was undertaken, as subjects were asked to rate the water features' displays as *manmade*, *natural* or *neither*.

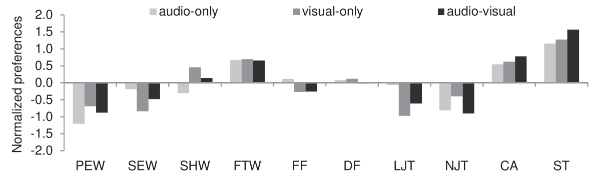

FIG. 3. Preferred water features for audio-only, visual-only, and audio-visual tests: Normalized preference values as a function of water features (refer to Table I for definitions of acronyms).

C. Participants

Here, 44 subjects who reported normal hearing ability took part in the three tests (21 males and 23 females) and 38 of those passed the consistency test (consistent judgments within a 95% confidence interval) and were retained for the analysis of results (19 males and 19 females). The age distribution of the retained subjects ranged from 24 to 47 yr (mean 30.1 yr and standard deviation 4.47 yr), and the cultural groups were composed of 19 "White" subjects, 14 "Middle Eastern," four "Asian," and one "African." As the sample size of each cultural group was small, cultural variations were however not analyzed statistically.

III. AUDIO-VISUAL PREFERENCES AND INTERACTION

A. Results and analysis

1. Audio-only, visual-only, and audio-visual preferences

Preferences obtained from the three paired comparison tests are shown as a bar chart in Fig. 3 and in ranked order in Table II [preferences defined over the range −2 (never preferred) to +2 (always preferred)]. Audio-only preferences show that the preferred water sounds were ST, FTW, and CA. Conversely, the least liked water sounds were SHW, NJT, and PEW. These preference results were significantly correlated with water features' objective categories (Spearman test, $p < 0.01$ for LJT = category 2, and $p < 0.05$ for LJT = category 3), confirming the findings of Galbrun and Ali:[7] Natural streams tend to be preferred to fountains, which are in turn preferred to waterfalls.

The preferred water displays obtained from the visual-only test were ST, FTW, and CA, results that are identical to those found for audio-only preferences. The least liked water displays were PEW, SEW, and LJT. These results also show that the single upward jets (LJT and NJT) tended not to be liked, unlike multiple upward jets that were identified as visually pleasing by previous research.[13] Correlations between these visual-only preferences and water features' objective categories were not statistically significant (Spearman test, $p > 0.05$).

Finally, preferred water features obtained from the audio-visual test were ST, CA, and FTW. Again, these results are very similar to those found for audio-only and visual-only preferences. Furthermore, the least liked water features were LJT, PEW, and NJT. Correlations between these audio-visual preferences and water features' objective categories were not statistically significant (Spearman test, $p > 0.05$).

The results of Fig. 3 and Table II indicate a positive effect of the visual stimulus on audio-visual preferences for

TABLE II. Ranking of preferred water features obtained from the audio-only, visual-only and audio-visual tests (refer to Table I for definitions of acronyms). The preferences are listed as normalized preference values.

	Audio-only		Visual-only		Audio-visual	
Ranking position	Sound code	Norm. pref.	Sound code	Norm. pref.	Sound code	Norm. pref.
1	ST	1.16	ST	1.27	ST	1.57
2	FTW	0.67	FTW	0.70	CA	0.78
3	CA	0.55	CA	0.62	FTW	0.65
4	FF	0.12	SHW	0.46	SHW	0.14
5	DF	0.08	DF	0.12	DF	−0.02
6	LJT	−0.07	FF	−0.27	FF	−0.26
7	SEW	−0.19	NJT	−0.40	SEW	−0.48
8	SHW	−0.30	PEW	−0.69	LJT	−0.61
9	NJT	−0.81	SEW	−0.84	PEW	−0.88
10	PEW	−1.20	LJT	−0.97	NJT	−0.90

ST, CA, SHW, and PEW (i.e., audio-visual scores are higher than audio-only scores). However, the presentation of visual material negatively affected audio-visual preferences for SEW, FF, DF, LJT, NJT, and (very marginally) FTW. According to an independent sample t-test, the comparison between audio-only and audio-visual preferences indicated that mean differences in preference scores are significant only for ST [$t(74) = -2.53$, $p < 0.05$], the visual stimulus significantly increasing preference scores [mean normalized preference scores of 1.16 (SD = 0.86) and 1.57 (SD = 0.51) for the audio-only and audio-visual tests, respectively]. Similarly, the comparison between visual-only and audio-visual preferences showed significant mean differences only for NJT [$t(74) = 2.27$, $p < 0.05$], the auditory stimulus significantly decreasing preference scores [mean normalized preference scores of −0.40 (SD = 1.02) and −0.90 (SD = 0.91) for the visual-only and audio-visual tests, respectively]. These t-test results suggest that an added stimulus (either visual or auditory) only rarely leads to a significant change in preferences.

Statistically significant correlations were found between the average ranking positions of the three tests (Spearman test): $\rho = 0.71$ with $p < 0.05$ for audio-only vs visual-only, $\rho = 0.83$ with $p < 0.01$ for audio-only vs audio-visual, and $\rho = 0.76$ with $p < 0.05$ for visual-only vs audio-only. Additional correlation analysis was carried out using the preference data of all subjects (Table III), rather than averages, and results showed no correlations between audio-only and visual-only preferences (although the significant correlations obtained using averages are important and should not be discarded, as the use of average values decreases the errors within the test). However, Table III shows that statistically significant correlations were found between audio-only and audio-visual preferences for seven out of ten water features (SHW, PEW, FF, DF, FTW, LJT, and NJT), and between visual-only and audio-visual preferences again for seven out of ten water features (CA, ST, SHW, PEW, FF, DF, FTW). Furthermore, correlations occurred between both uni-modal scores and audio-visual scores for five out of ten features (SHW, PEW, FF, DF, FTW), while a single stimulus appeared to be dominant in the audio-visual rating only for a minority of features (visual dominance for CA and ST,

TABLE III. Correlations (correlation coefficient ρ, Spearman test) between audio-only, visual-only and audio-visual preferences, for ten different water features used over road traffic noise (refer to Table I for definitions of acronyms).[a]

Sound code	Audio-only vs Visual-only	Audio-only vs Audio-visual	Visual-only vs Audio-visual
CA	−0.12	0.19	0.53**
ST	−0.02	0.29	0.47**
SEW	−0.32	0.24	0.17
SHW	0.28	0.49**	0.35*
PEW	0.26	0.32*	0.37*
FF	0.00	0.41*	0.52**
DF	0.03	0.37*	0.39*
FTW	0.05	0.33*	0.48**
LJT	−0.05	0.56**	0.03
NJT	0.08	0.63**	0.25

[a]* Significant correlation at the 0.05 level ($p < 0.05$); ** Significant correlation at the 0.01 level ($p < 0.01$).

and auditory dominance for LJT and NJT), as already pointed out by the t-test results.

No statistically significant differences in responses were found between different gender and age groups in both uni-modal and bi-modal preference tests (Mann-Whitney test, $p > 0.05$).

2. Hierarchical cluster analysis of preferences

The low coefficients of concordance obtained from the preference results (Kendall's coefficient of concordance $W_{\text{audio-only}} = 0.33$, $W_{\text{visual-only}} = 0.35$, and $W_{\text{audio-visual}} = 0.41$)[26,27] justified a hierarchical cluster analysis, in view of identifying more consistent groups of subjects. The criteria applied for the analysis were the average linkage method and the Square Euclidian distance.[26] Results showed that subjects can be split into two clusters for audio-only and audio-visual tests, and into three clusters for the visual-only test, all clusters showing an increase in Kendall's coefficient of concordance (Tables IV and V). For the audio-only preference clusters, the sounds most affected by the analysis were LJT (variation of up to 8 ranking positions between clusters) and, to a lower degree, SHW (±5 positions). Similar results were obtained for audio-visual preferences, as the features the most affected by cluster analysis were again LJT (±8 positions) and SHW (±6 positions). This indicates that LJT and SHW were either liked or disliked, and this separated them into two distinct preference groups. For the visual-only preference clusters, the water features most affected by the analysis were FTW (±7 positions), NJT (±4 positions), PEW (±6 positions), and SEW (±8 positions). However, consistent visual-only preferences were found between most of the subjects (31 out of the 38 subjects in cluster 1).

3. Correlations between acoustic/psychoacoustic parameters and audio-only preferences

Previous research by Galbrun and Ali[7] suggested that no acoustical and psychoacoustical parameter correlated well with individual sound preferences, but analysis made on ranked groups of sounds indicated that low sharpness and large temporal variations ($L_{A10} - L_{A90}$) were preferred on

TABLE IV. Ranking of preferences obtained for (a) the audio-only test and (b) the audio-visual test from all subjects retained for the analysis and from clusters obtained from hierarchical cluster analysis. The preferences are listed as normalized preference values. Kendall's coefficient of concordance, W, is also given for results including all subjects and for the clusters.

	All subjects		Cluster 1 (17 subjects)		Cluster 2 (21 subjects)	
Ranking	Sound code	Norm. pref.	Sound code	Norm. pref.	Sound code	Norm. pref.
(a)						
1	ST	1.16	CA	0.80	ST	1.53
2	FTW	0.67	FTW	0.72	LJT	1.15
3	CA	0.55	ST	0.69	FTW	0.62
4	FF	0.12	SHW	0.69	FF	0.37
5	DF	0.08	DF	0.59	CA	0.35
6	LJT	−0.07	SEW	0.17	DF	−0.33
7	SEW	−0.19	FF	−0.20	NJT	−0.41
8	SHW	−0.30	PEW	−0.59	SEW	−0.48
9	NJT	−0.81	NJT	−1.29	SHW	−1.11
10	PEW	−1.20	LJT	−1.58	PEW	−1.70
W	0.33		0.51		0.66	
(b)						
1	ST	1.57	ST	1.56	ST	1.57
2	CA	0.78	LJT	1.23	CA	0.90
3	FTW	0.65	FTW	0.83	SHW	0.60
4	SHW	0.14	CA	0.51	FTW	0.58
5	DF	−0.02	FF	0.06	DF	0.09
6	FF	−0.26	DF	−0.30	SEW	−0.35
7	SEW	−0.48	NJT	−0.30	FF	−0.39
8	LJT	−0.61	SEW	−0.79	PEW	−0.50
9	PEW	−0.88	SHW	−0.99	NJT	−1.14
10	NJT	−0.90	PEW	−1.80	LJT	−1.36
W	0.41		0.71		0.54	

average.[7] Correlations between acoustical/psychoacoustical parameters and audio-only preferences obtained here (Table II) are shown in Table VI (both for all subjects and for clusters). No statistically significant correlations were found with sound preferences including all subjects (confirming previous work[7]). Preferences from all subjects tended to be related to higher temporal variation (positive ρ), as well as lower sharpness (negative ρ), but the correlations were not statistically significant (Spearman test, $p > 0.05$). Furthermore, analysis made on ranked groups of sounds also showed no significant correlations, unlike the previous results obtained by Galbrun and Ali.[7] For sharpness, this is probably due to the absence of flat and hard impact materials in the current research, as these can significantly increase sharpness and tend not to be liked.[7] In Ref. 7, hard materials increased the average sharpness calculated from low ranked sounds, thus providing significant correlations with sharpness for the analysis made on groups of sounds. However, results obtained for clusters showed some statistically significant correlations for temporal variation and roughness (cluster 2), as well as for pitch strength (cluster 1). But overall, current and previous results[7] suggest that there is a weak association between acoustical/psychoacoustical parameters and preferences of water sounds.

B. Main findings of paired comparisons

Statistically significant correlations occurred between audio-only preferences and objective categories of water features. These indicated that natural streams tend to be preferred to fountains, which are in turn preferred to waterfalls, confirming previous work.[7]

Both auditory and visual stimuli tended to affect preferences. Auditory and visual stimuli were equally important in the audio-visual assessment of half of the water features tested, with one stimulus being dominant only in a minority of features. This reflects the interdependence between unimodal and bi-modal perception and suggests that equal attention should be given to the design of both stimuli.

The addition of a visual stimulus increased preferences in some cases (four out of ten water features), but decreased them in other cases (six out of ten features). As paired comparisons were used, an increase in preference scores for some features necessarily led to a decrease for other features. Therefore, these results do not mean that some visual stimuli are detrimental, but simply that some features benefit more than others from a visual stimulus. Rating scales used in waterscapes studies[2,6] showed that the addition of a visual stimulus improves perception most of the time (compared to audio-only perception).

Although preference scores changed when a stimulus was added, mean differences indicated that an added

TABLE V. Like Tables IV(a) and IV(b), but for the visual-only test.

	All subjects		Cluster 1 (31 subjects)		Cluster 2 (4 subjects)		Cluster 3 (3 subjects)	
Ranking	Sound code	Norm. pref.	Sound code	Norm. pref.	Sound code	Norm. pref.	Sound code	Norm. pref.
1	ST	1.27	ST	1.34	ST	1.67	CA	1.70
2	FTW	0.70	FTW	1.01	SEW	1.00	PEW	1.56
3	CA	0.62	CA	0.51	NJT	0.78	SHW	1.11
4	SHW	0.46	SHW	0.42	CA	0.67	ST	0.07
5	DF	0.12	DF	0.19	SHW	0.22	DF	−0.22
6	FF	−0.27	FF	−0.15	DF	−0.22	SEW	−0.37
7	NJT	−0.40	NJT	−0.49	LJT	−0.67	FTW	−0.37
8	PEW	−0.69	PEW	−0.78	FF	−0.89	FF	−0.67
9	SEW	−0.84	LJT	−0.92	FTW	−0.89	NJT	−0.96
10	LJT	−0.97	SEW	−1.13	PEW	−1.67	LJT	−1.85
W	0.35		0.43		0.73		0.82	

TABLE VI. Correlation coefficient (ρ, Spearman test) showing correlations between acoustic/psychoacoustic parameters and audio-only preferences of water sounds in the presence of road traffic noise. Results are presented for preferences obtained from all subjects retained for the analysis, and for preferences from the clusters obtained from hierarchical cluster analysis.[a]

	$L_{A10}-L_{A90}$	$L_{Ceq}-L_{Aeq}$	Sharpness	Roughness	Pitch strength
Audio-only preferences—All subjects	0.23	−0.07	−0.15	0.51	0.35
Audio-only preferences—Cluster 1	−0.46	−0.40	0.52	−0.15	0.64*
Audio-only preferences—Cluster 2	0.64*	0.10	−0.50	0.82**	0.14

[a]* Significant correlation at the 0.05 level ($p<0.05$); ** Significant correlation at the 0.01 level ($p<0.01$).

stimulus (either visual or auditory) only rarely leads to a statistically significant change. This suggests that a single stimulus is rarely dominant in driving waterscapes' preferences (as already indicated by correlations between uni-modal and bi-modal perception).

Hierarchical cluster analysis of audio-only and audio-visual preferences indicated that some water features can be either clearly liked or disliked by different subjects, although this tends to be unusual (observed here for two features out of ten).

Current and previous research[7] suggest that there is a weak association between acoustical/psychoacoustical parameters and preferences of water sounds.

IV. SEMANTIC ANALYSIS, CATEGORIZATION, AND EVOCATION

A. Results and analysis

1. Semantic characterization of water sounds and correlations between components

The semantic characterization of each water sound is shown in Fig. 4 (the components shown on the right are discussed in Sec. IV A 2). As mentioned in Sec. II B, a five-point verbal scale was used for the tests, and the results presented correspond to averages taken across all subjects (range −2 to +2). Overall, Fig. 4 shows that positive values tended to be obtained for relaxation, naturalness, freshness, and familiarity, while results were more scattered between negative and positive values for the remaining attributes. This was analyzed in more detail by looking at correlations (Spearman test) between the average scores of attributes (Table VII), which showed that relaxation, naturalness, freshness, and familiarity were significantly correlated with each other ($p<0.01$, except for the correlation between freshness and familiarity where $p<0.05$). A significant correlation was also found between the perceived sharpness and the perceived roughness ($p<0.01$), and these attributes also showed significant negative correlations with relaxation, freshness, naturalness and familiarity ($p<0.01$, except for the correlations between sharpness and relaxation where $p<0.05$). This indicates that high values of perceived sharpness or perceived roughness were associated to low ratings of relaxation, naturalness, freshness, and familiarity. Furthermore, envelopment was positively correlated with speed ($p<0.01$) and negatively correlated with temporal variation ($p<0.01$). This suggests that water sounds perceived as more enveloping tended to have higher flow rates and were steadier.

No statistically significant differences in responses were found between different gender and age groups (Mann-Whitney test, $p>0.05$), with only few exceptions observed in the evaluation of freshness and familiarity for different genders.

2. Principal component analysis of semantic components and correlations with preferences

A principal component analysis (PCA) was carried out in view of grouping the nine semantic attributes into a lower number of components. The PCA was based on the varimax rotation method to extract the orthogonal components, and the criterion of eigenvalues greater than 1 was applied.[26] Results given in Table VIII indicate that three main components were found to be important for waterscapes' characterization. The first component included qualitative properties of the water sounds and was called "emotional assessment." The second and third components were related to psychoacoustical and acoustical properties of sounds and were called "sound quality" and "envelopment and temporal variation," respectively. Component 1 explained 32% of the total variance, followed by component 2 with 20%, and component 3 with 14%.

Table VIII indicates that correlations (Spearman test) between audio-only preferences and semantic components (average scores) showed a statistically significant positive correlation for component 1 ($p<0.01$) and a statistically significant negative correlation for component 2 ($p<0.01$). This suggests that the "emotional assessment" positively contributed to audio-only preferences of the water sounds, while high values of "sound quality" negatively affected these preferences. However, the negative correlation obtained for the "sound quality" component does not mean poor sound quality: Negative values are simply due to the poor ratings of high perceived sharpness, perceived roughness, and speed. Furthermore, Table VIII shows that correlations with individual semantic attributes (average scores) indicated statistically significant correlations between audio-only preferences and relaxation ($p<0.01$), naturalness ($p<0.05$), freshness ($p<0.01$), perceived sharpness ($p<0.05$), and perceived roughness ($p<0.01$). These results suggest that the subjective perception of waterscapes depended mainly on the emotional attributes associated to the sound (component 1) and the characterization of its sound quality (component 2). Conversely, envelopment and temporal variation (component 3) had no significant impact toward audio-only preferences.

Logistic regression[26,28] was also examined to determine the stochastic relationship between audio-only preferences and semantic components/attributes. Results presented here are limited to binary logistic regression and the relation with components, as multinomial logistic regression and models found for attributes showed an unacceptable level of accuracy. Audio-only preferences were defined in the model as the dependent variable (the objective being to predict the probability of preferences), while semantic components were the independent variables. In binary logistic regression, the dependent variable used for identifying the model must be dichotomous.[28] In the current case, the binary coding

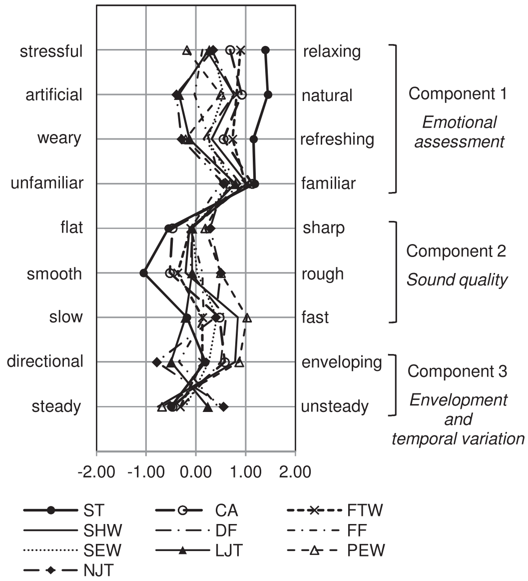

FIG. 4. Semantic characterization of each water sound used over road traffic noise, illustrating both attributes and components. Results are given as average scores obtained for each attribute (refer to Table I for definitions of acronyms).

used for each subject's preferences (9 paired comparisons for each water sound) was as follows: 0 = sound preferred 0–4 times (low level of satisfaction); 1 = sound preferred 5–9 times (high level of satisfaction). The number of subjects in each binary category was then calculated for each water sound, the greater number defining the preference category (0 or 1). The components' data was calculated by averaging the ratings obtained from all attributes included in each component (data shown in Fig. 4). The binary categories and components' values were used as input data to determine the coefficients α and β of the following logit model

$$\begin{aligned}\mathrm{PREF}_{\mathrm{Audio\text{-}only}} = \alpha &+ \beta_{\mathrm{Emot\ Assess}}\,\mathrm{EMOT\ ASSESS}\\ &+ \beta_{\mathrm{Sound\ Qual}}\,\mathrm{SOUND\ QUAL}\\ &+ \beta_{\mathrm{Env\ \&\ Temp\ Var}}\,\mathrm{ENV\ \&\ TEMP\ VAR},\end{aligned} \quad (1)$$

where $\mathrm{PREF}_{\mathrm{Audio\text{-}only}}$ is the dependent variable calculated by the logit model (range $-\infty$ to $+\infty$), and EMOT ASSESS, SOUND QUAL, and ENV & TEMP VAR are the independent variables (average ratings for each component, in the range -2 to $+2$). The modeling data obtained for the logit model of Eq. (1) is given in Table IX, which shows a 66.3% accuracy of the predicting model and a Nagelkerke R^2 value (analogous of the R^2 value applied in linear regression)[26] of 0.17 (reasonable fit). Furthermore, the likelihood ratio test based on the $-2LL$ ratio[28] was significant at the 0.05 level, meaning that the model with all predictors (independent variables) was significantly different from the one including only the constant α. The most important finding of this logit model is that EMOT ASSESS is the only independent variable to be a statistically significant predictor for the model ($p < 0.01$), confirming that "emotional assessment" explains most preferences. The positive sign found for $\beta_{\mathrm{Emot\ Assess}}$ in Table IX also indicates that the likelihood in giving a positive audio-only rating increases as the rating of "emotional assessment" increases.

To understand the usefulness of the model given by Eq. (1), it is important to note that the outcome of logistic regression is not a prediction of the dependent variable's value ($\mathrm{PREF}_{\mathrm{Audio\text{-}only}}$), as in the linear regression, but a probability, $P(\mathrm{PREF}_{\mathrm{Audio\text{-}only}})$, of belonging to one of the two categories used as the input data for $\mathrm{PREF}_{\mathrm{Audio\text{-}only}}$.[28] In this case, it is the probability of a high level of satisfaction (category 1) from the water sound in terms of relaxation and peacefulness, and the probability can be expressed as[28]

$$P(\mathrm{PREF}_{\mathrm{Audio\text{-}only}}) = 1/[1 + e^{-(\mathrm{PREF}_{\mathrm{Audio\text{-}only}})}]. \quad (2)$$

Table X shows the probability results obtained for the logit model, as a function of the actual *EMOT ASSESS* values and the $\mathrm{PREF}_{\mathrm{Audio\text{-}only}}$ values calculated from Eq. (1), with $\beta_{\mathrm{Sound\ Qual}} = \beta_{\mathrm{Env\ \&\ Temp\ Var}} = 0$ (components ignored because not statistically significant in the model). The table indicates that the probability of belonging to category 1 (high level of satisfaction) is at least 70% for an "emotional assessment" rating of +1 or more (on a -2 to $+2$ range). Table IX also shows an odd-ratio[26] of 3.25 for EMOT ASSESS, meaning that people who have positively rated the waterscapes in terms of "emotional assessment" (range 0 to +2), are 3.25 times more likely to give high preference scores to waterscapes in the audio-only condition.

A multiple linear regression analysis[26] was also run using the three principal components as dependent variables and provided exactly the same findings (i.e., significant relation found only between "emotional assessment" and audio-only preferences), with $R^2 = 0.16$, $F(3, 376) = 24.31$ and

TABLE VII. Correlations between semantic attributes (correlation coefficient ρ, Spearman test).[a]

	Relaxation	Naturalness	Freshness	Familiarity	Perceived sharpness	Perceived roughness	Speed	Envelopment	Temporal variation
Relaxation	1.00	0.84**	0.81**	0.78**	−0.69*	−0.87**	−0.38	0.09	−0.30
Naturalness	0.84**	1.00	0.86**	0.88**	−0.80**	−0.89**	0.05	0.48	−0.67*
Freshness	0.81**	0.86**	1.00	0.74*	−0.93**	−0.87**	−0.27	0.16	−0.45
Familiarity	0.78**	0.88**	0.74*	1.00	−0.77**	−0.94**	−0.06	0.37	−0.66*
Perceived sharpness	−0.69*	−0.80**	−0.93**	−0.77**	1.00	0.87**	0.20	−0.22	0.48
Perceived roughness	−0.87**	−0.89**	−0.87**	−0.94**	0.87**	1.00	0.25	−0.19	0.49
Speed	−0.38	0.05	−0.27	−0.06	0.20	0.25	1.00	0.81**	−0.49
Envelopment	0.09	0.48	0.16	0.37	−0.22	−0.19	0.81**	1.00	−0.84**
Temporal variation	−0.30	−0.67*	−0.45	−0.66*	0.48	0.49	−0.49	−0.84**	1.00

[a]* Significant correlation at the 0.05 level ($p < 0.05$); ** Significant correlation at the 0.01 level ($p < 0.01$).

TABLE VIII. Correlations (correlation coefficient ρ, Spearman test) between audio-only preferences and components obtained from principal component analysis, as well as between audio-only preferences and semantic attributes.[a]

Component	Correlation coefficient (ρ)	Attribute	Correlation coefficient (ρ)
1—Emotional assessment	0.82**	Relaxation	0.83**
		Naturalness	0.69*
		Familiarity	0.57
		Freshness	0.83**
2—Sound quality	−0.88**	Perceived sharpness	−0.75*
		Perceived roughness	−0.79**
		Speed	−0.57
3—Envelopment and temporal variation	−0.30	Temporal variation	−0.4
		Envelopment	−0.20

[a]* Significant correlation at the 0.05 level ($p < 0.05$); ** Significant correlation at the 0.01 level ($p < 0.01$).

$p = 0.000$ (95% confidence intervals of 3.74 and 4.27 for the model's constant (lower and upper bounds), and of 1.10 and 1.85 for the coefficient of "emotional assessment").

Correlations (Spearman test) between semantic components (average scores) and acoustic/psychoacoustic parameters (Table XI) indicated statistically significant correlations between $L_{A10} - L_{A90}$ and component 3 ($p < 0.05$), between roughness and component 2 as well as component 3 ($p < 0.05$), and between pitch strength and component 1 ($p < 0.05$). A further analysis of correlations with semantic attributes clarified that $L_{A10} - L_{A90}$, sharpness and roughness were significantly correlated with speed and envelopment (Table XII). Furthermore, pitch strength was also significantly correlated with familiarity (Table XII). These results indicate that subjects were unable to correctly assess the sharpness, roughness, and temporal variations of the water sounds, as no correlations were found between physical parameters and their corresponding perceptual descriptors. Conversely, the perception of speed and envelopment were strongly correlated with acoustic ($L_{A10} - L_{A90}$) and psychoacoustic (sharpness and roughness) parameters, but no reasons could be found to justify these correlations.

Finally, it is interesting to note that the natural stream (ST) was poorly rated in terms of envelopment: This was not expected due to the strong spatial quality reflected in the left and right channels of its binaural recording, as this sound was measured at the junction of two streams. This was probably be due to people rating envelopment as a quality for which no direction can be associated to the sound, rather than a well-defined stereo field.

TABLE IX. Logit model data for predicting audio-only preferences in relation to semantic components. The model fitting information.

Accuracy of predicting the model		66.3%	
Nagelkerke R^2 (range 0–1)		0.17	
Predictors	Coefficients (β; α)	p-value	Odds-ratio
EMOT ASSES	1.18	0.00[a]	3.25[b]
SOUND QUAL	0.18	0.38	-
ENV & TEMP VAR	0.02	0.89	-
Constant α	−0.32	0.01	0.73

[a]Significant correlation at the 0.01 level ($p < 0.01$).
[b]Increase in odds of positively rating water sounds if subjects positively rate the "emotional assessment."

TABLE X. Logit model data for predicting audio-only preferences in relation to semantic components. The probability of high levels of satisfaction based on Eq. (1), with $\beta_{Sound\,Qual} = \beta_{Env\,and\,Temp\,Var} = 0$ (i.e., "emotional assessment" component used as the only predictor).

EMOT ASSESS [−2;+2]	$PREF_{Audio\text{-}only}$ [$-\infty;+\infty$]	Probability of high levels of satisfaction [0;1]
−2	−2.68	0.06
−1	−1.50	0.18
0	−0.32	0.42
1	0.86	0.70
2	2.04	0.88

3. Evocation and qualitative categorization

Figure 5 illustrates the results obtained for the manmade evocation of water sounds used over road traffic noise. The sounds with the highest manmade evocation were the single jets (NJT and LJT), which were qualitatively described in open-ended questions as water tap sounds (Table XIII). The foam fountain (FF) was also identified as a manmade sound, and qualitatively described by several subjects as a washing sound (Table XIII). All the other water sounds were considered by the majority of subjects as not manmade, especially the plain edge waterfall (PEW) which was not manmade for 87% of the subjects but was nevertheless not liked [sound described as a waterfall (Table XIII)]. A negative correlation was found between audio-only preferences and manmade evocation, but it was not statistically significant (Spearman test, $\rho = -0.049$, $p > 0.05$).

Subjects were also asked to indicate whether the water sounds made them think about rainfall. Figure 6 indicates that the small holes' edge waterfall (SHW) and the dome fountain (DF) evoked rainfall to around 80% of the subjects, and these sounds were also qualitatively described by most subjects as rainfall (Table XIII). The plain edge waterfall (PEW) and the fountain with multiple upward jets (FTW) also resembled rainfall for a majority of subjects. Conversely, the single jets (NJT and LJT), as well as the

TABLE XI. Correlations (correlation coefficient ρ, Spearman test) between semantic components and acoustic/psychoacoustic parameters. Correlations calculated from sounds including both road traffic noise and water sounds.

Parameter	Component 1 Emotional assessment	Component 2 Sound quality	Component 3 Envelopment and temporal variation
L_{A10}–L_{A90}	0.70	0.45	0.75[a]
L_{Ceq}–L_{Aeq}	0.45	0.13	−0.20
Sharpness	−0.18	−0.33	−0.50
Roughness	−0.17	0.67[a]	0.71[a]
Pitch strength	−0.64[a]	0.14	0.24

[a]Significant correlation at the 0.05 level ($p < 0.05$).

TABLE XII. Correlations (correlation coefficient ρ, Spearman test) between semantic attributes and acoustic/psychoacoustic parameters. Correlations calculated from sounds including both road traffic noise and water sounds.[a]

Parameter	Relaxation	Naturalness	Familiarity	Freshness	Perceived sharpness	Perceived roughness	Speed	Envelopment	Temporal variation
L_{A10}-L_{A90}	−0.12	0.31	0.17	0.01	−0.26	−0.05	0.88**	0.79**	−0.55
L_{Ceq}-L_{Aeq}	0.17	0.39	0.42	0.22	−0.20	−0.34	0.22	0.11	−0.21
Sharpness	0.11	−0.37	−0.33	0.51	0.18	0.20	−0.76**	−0.68*	0.59
Roughness	−0.26	0.03	−0.12	0.25	0.22	0.24	0.85**	0.74*	−0.47
Pitch strength	−0.35	−0.57	−0.64*	0.21	0.39	0.60	−0.07	−0.07	0.28

[a]* Significant correlation at the 0.05 level ($p < 0.05$). ** Significant correlation at the 0.01 level ($p < 0.01$).

foam fountain (FF) and natural stream (ST), did not resemble rainfall for more than 80% of the subjects. A negative correlation was found between audio-only preferences and rainfall evocation, but it was not statistically significant (Spearman test, $\rho = -0.159$, $p > 0.05$).

In addition to water sounds' perception questions, subjects were also asked to indicate whether the water features' displays (Fig. 1) looked natural, manmade, or neither. Figure 7 shows that the most naturally looking water feature was the natural stream (ST), followed by the small holes' edge waterfall (SHW) and the plain edge waterfall (PEW). All the other water features were categorized as manmade looking by the majority of subjects, with percentages above 70% obtained for FF, FTW, NJT, and SEW. No statistically significant correlations were found between this visual categorization and the visual-only and audio-visual preferences (Spearman test, $p > 0.05$), although all correlations between preferences and natural scores were positive, while all correlations between preferences and manmade scores were negative. These results also explain the increase in audio-visual preference scores (compared to audio-only scores) found for ST, SHW, and PEW (see Sec. III A 1): Natural looking features tend to affect perception positively, although a manmade looking feature such as the cascade (CA) also improved perception (i.e., manmade looking features can be visually pleasing).

In view of categorizing water features, subjects were asked to indicate whether the water sounds made them think of a waterfall, a fountain, a natural stream or none of these. Results are given as perceived categories in Table XIV, together with the objective categories previously defined in Table I. It can be noted that LJT is listed there both as objective category 2 (fountain) and category 3 (stream) (refer to the justification given in Sec. II). Significant correlations were found between the objective and perceived categories when LJT was objectively classified as a fountain (Spearman test, $p < 0.05$). Subjects were however unable to categorize most of the water sounds, as most of the highest percentages occurred in perceived category 4 (*none of these*). An exception was represented by the natural stream (ST) and the cascade (CA), which were clearly identified by subjects as natural stream sounds. The latter was also confirmed by their qualitative description (Table XIII). This suggests that the categorization into waterfalls and fountains is difficult, but subjects not selecting *none of these* [around 60% of subjects on average (excluding ST and CA responses)] tended to accurately identify the type of water feature that they heard. It can also be noted that, on average, waterfalls were more identifiable than fountains.

In the qualitative assessment of water sounds, subjects could also write down what the sound made them think of (if anything), and the most commonly mentioned qualities are listed in Table XIII. These descriptions confirm the findings previously discussed in terms of manmade sounds (NJT, LJT, and FF), rainfall (SHW and DF), natural streams (ST and CA), and waterfalls (PEW). Additionally, it can be noted that the sawtooth edge waterfall (SEW) was described as containing multiple sources. SEW was indeed made of multiple streams, i.e., multiple impact areas/sources. Furthermore, the fountain with multiple upward jets (FTW) was evocative of a courtyard's water feature for some subjects. This was recurrently mentioned by Middle Eastern subjects who are familiar with courtyard architecture including water features, suggesting that cultural factors might affect the evocation of water sounds.

Finally, it can be noted that for all the tests presented in this section, no statistically significant differences in responses were found between different gender and age

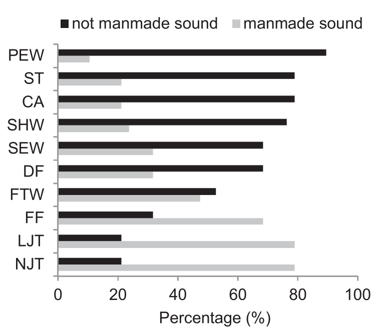

FIG. 5. Manmade sound evocation for the ten water sounds tested in the presence of road traffic noise (refer to Table I for definitions of acronyms).

TABLE XIII. Qualitative "open-ended" descriptions of water sounds in the presence of road traffic noise (refer to Table I for the definitions of acronyms). The qualities listed correspond to those most commonly mentioned.

Sound code	Qualitative description of the sound
PEW	Waterfall
SEW	Multiple sources (water and noise)
SHW	Rainfall
FTW	Water feature in courtyard
DF	Rainfall
FF	Washing, manmade sound
LJT	Tap, manmade sound
NJT	Tap, manmade sound
CA	Natural stream
ST	Natural stream

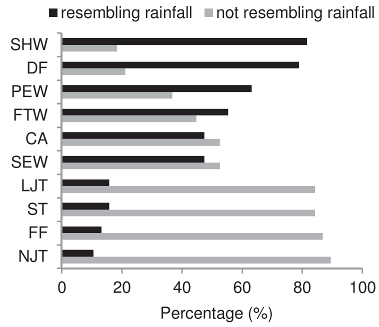

FIG. 6. Rainfall evocation for the ten water sounds tested in the presence of road traffic noise (refer to Table I for definitions of acronyms).

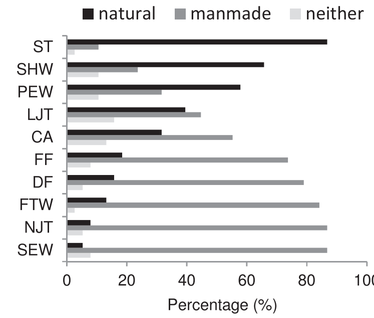

FIG. 7. Visual categorization (natural vs manmade) of the ten water features tested (refer to Table I for definitions of acronyms).

groups (Mann-Whitney test, $p > 0.05$), with only very few exceptions that are not noteworthy.

B. Main findings of semantic analysis, categorization, and evocation

Three main components have been found to be important for waterscapes' characterization: "emotional assessment" (component 1), "sound quality" (component 2), and "envelopment and temporal variation" (component 3). Only the first two components showed statistically significant correlations with audio-only preferences, component 1 being the most important, as it explained most preferences and was the only significant predictor of logistic regression.

For component 1, the attributes relaxation, naturalness and freshness showed statistically significant positive correlations with audio-only preferences, while for component 2, perceived sharpness and perceived roughness showed significant negative correlations.

Subjects were unable to correctly assess the sharpness, roughness and temporal variations of water sounds, as no correlations were found between physical parameters and their corresponding perceptual descriptors.

Single upward jets were perceived by most subjects as manmade sounds evocative of water taps. Furthermore, manmade sounds tended not to be preferred (negative correlation with audio-only preferences), but the negative correlation was not statistically significant.

Four out of the ten water sounds tested were evocative of rainfall. The correlation between rainfall evocation and audio-only preferences was negative, but not statistically significant.

Natural looking features tended to be preferred in visual-only and audio-visual tests, but correlations with preferences were not statistically significant. Furthermore, natural looking features tended to increase audio-visual preference scores compared to audio-only scores, while manmade looking features tended to decrease them (although the mean differences were statistically significant only for ST, as pointed out in Sec. III A 1).

Natural stream sounds were easily identifiable, unlike waterfall and fountain sounds.

V. CONCLUSIONS

This research examined the audio-visual interaction of water features used over road traffic noise, as well as their semantic aural assessment, perceptual categorization, and evocation properties. The work had the objective of characterizing water sounds, rather than examining their masking ability. All the tests were carried out within the context of peacefulness and relaxation, with the aim of being able to provide evidence-based design solutions for small to medium sized water features that have applications for road traffic noise masking.

TABLE XIV. Percentages of perceived water features' categories and correlation coefficient (ρ, Spearman test) showing correlations between the objective categories and the perceived categories (1 = waterfall; 2 = fountain; 3 = natural stream; 4 = none of these). Tests carried out with sounds including both water sounds and road traffic noise.

Water Feature	Objective category	Perceived category 1 (%)	Perceived category 2 (%)	Perceived category 3 (%)	Perceived category 4 (%)
PEW	1—Waterfall	42.1	0	7.9	50
SEW	1—Waterfall	23.7	18.4	39.5	18.4
SHW	1—Waterfall	28.9	10.5	13.2	47.4
FTW	2—Fountain	18.4	23.7	23.7	34.2
DF	2—Fountain	34.2	10.5	21.1	34.2
FF	2—Fountain	5.3	34.2	23.7	36.8
LJT	2/3—Fountain/Stream	10.5	28.9	13.2	47.4
NJT	2—Fountain	10.5	36.8	7.9	44.7
CA	3—Stream	18.4	5.3	65.8	10.5
ST	3—Stream	2.6	13.2	78.9	5.3
Correlation coefficient (ρ)		0.65[a]	0.70[a]/ 0.53	0.70[a]/ 0.46	-

[a]Correlation is significant at the 0.05 level ($p < 0.05$). Perceived category 1 = waterfall; Perceived category 2 = fountain; Perceived category 3 = natural stream; Perceived category 4 = none of these. Two correlation coefficients given for perceived categories 2 and 3: the value on the left corresponds to LJT being assigned objective category 2, whilst the value on the right corresponds to LJT being assigned objective category 3.

Paired comparisons highlighted the inter-dependence between uni-modal (audio-only or visual-only) and bi-modal (audio-visual) perception, suggesting that equal attention should be given to the design of both stimuli. Audio-visual analysis also indicated that natural looking water features tend to increase preferences (compared to audio-only preferences), while manmade looking features tend to decrease preferences (although these paired comparisons' results do not mean that the addition of a visual stimulus is detrimental for manmade looking features). Furthermore, it was found that an added stimulus (either visual or auditory) only rarely leads to a statistically significant change in preference scores, thus suggesting that a single stimulus is rarely dominant in driving preferences.

Results obtained from audio-only paired comparisons confirmed previous research,[7] showing that natural streams tend to be preferred to fountains, which are in turn preferred to waterfalls. Analysis of correlations indicated that the association between acoustical/psychoacoustical parameters and preferences of water sounds is weak, while semantic attributes and components showed significant correlations with preferences and therefore appear to be more reliable design descriptors compared to physical parameters. A principal component analysis identified the following three components within the semantic attributes tested: "Emotional assessment," "sound quality," and "envelopment and temporal variation." The first two showed statistically significant correlations with audio-only preferences, "emotional assessment" proving to be the most important component, as it explained most of the preference scores and was the only significant predictor identified by logistic regression. Within this component, naturalness, relaxation, and freshness showed significant correlations with audio-only preferences, while low scores of perceived sharpness and perceived roughness provided significant correlations within the "sound quality" component (although these perceived attributes were not correlated with the corresponding psychoacoustical descriptors).

Evocation results indicated that single upward jets are evaluated by most subjects as manmade sounds evocative of water taps, which probably justifies their low auditory rating. This was the only obvious association identified between evocation and preferences, as no significant correlations were found between preferences and manmade scores, rainfall scores, as well as familiarity scores. Overall, these results suggest that auditory evocation might be strictly associated to the overall perception of water features rather than to their auditory preferences.

Finally, categorization results indicated that natural stream sounds are easily identifiable, while it is difficult to identify waterfall and fountain sounds. Natural stream sounds could then be referred to as positive water soundmarks,[11] as they are easily recognizable and tend to be preferred for peacefulness and relaxation.

[1] J. Kang, *Urban Sound Environment* (Taylor and Francis, New York, 2007), pp. 98–99.

[2] G. Watts, R. Pheasant, K. Horoshenkov, and L. Ragonesi, "Measurement and subjective assessment of water generated sounds," Acta Acust. Acust. **95**, 1032–1039 (2009).

[3] J. Y. Jeon, P. J. Lee, J. You, and J. Kang, "Perceptual assessment of quality of urban soundscapes with combined noise sources and water sounds," J. Acoust. Soc. Am. **127**(3), 1357–1366 (2010).

[4] M. E. Nilsson, J. Alvarsson, M. Radsten-Ekman, and K. Bolin, "Auditory masking of wanted and unwanted sounds in a city park," Noise Control Eng. J. **58**, 524–531 (2010).

[5] B. De Coensel, S. Vanwetswinkel, and D. Botteldooren, "Effects of natural sounds on the perception of road traffic noise," J. Acoust. Soc. Am. **129**(4), EL148–EL153 (2011).

[6] J. Y. Jeon, P. J. Lee, J. You, and J. Kang, "Acoustical characteristics of water sounds for soundscape enhancement in urban open spaces," J. Acoust. Soc. Am. **131**(3), 2101–2109 (2012).

[7] L. Galbrun and T. T. Ali, "Acoustical and perceptual assessment of water sounds and their use over road traffic noise," J. Acoust. Soc. Am. **133**(1), 227–237 (2013).

[8] M. Rådsten-Ekman, Ö. Axelsson, and M. E. Nilsson, "Effects of sounds from water on perception of acoustic environments dominated by road traffic noise," Acta Acust. Acust. **99**, 218–225 (2013).

[9] J. Y. Hong and J. Y. Jeon, "Designing sound and visual components for enhancement of urban soundscapes," J. Acoust. Soc. Am. **134**(3), 2026–2036 (2013).

[10] M. Raimbault, C. Lavandier, and M. Berengier, "Ambient sound assessment of urban environments: Field studies in two French cities," Appl. Acoust. **64**(12), 1241–1256 (2003).

[11] R. M. Schafer, *Our Sonic Environment and the Soundscape: The Tuning of the World* (Destiny Books, Rochester, VT, 1994), 320 pp.

[12] J. You, P. J. Lee, and J. Y. Jeon, "Evaluating water sounds to improve the soundscape of urban areas affected by traffic noise," Noise Control Eng. J. **58**, 477–483 (2010).

[13] J. Nasar and Y-H. Lin, "Evaluative responses to five kinds of water features," Landsc. Res. **28**(4), 441–450 (2003).

[14] J. L. Carles, I. López Barrio, and J. V. De Lucio, "Sound influence on landscape values," Landsc. Urban Plan. **43**(4), 191–200 (1999).

[15] S. Viollon, C. Lavandier, and C. Drake, "Influence of visual setting on sound ratings in an urban environment," Appl. Acoust. **63**(5), 493–511 (2002).

[16] J. Carles, F. Bernáldez, and J. De Lucio, "Audio-visual interactions and soundscape preferences," Landsc. Res. **17**(2), 52–56 (1992).

[17] R. Pheasant, K. Horoshenkov, G. Watts, and B. Barrett, "The acoustic and visual factors influencing the construction of tranquil space in urban and rural environments tranquil spaces-quiet places?," J. Acoust. Soc. Am. **123**(3), 1446–1457 (2008).

[18] J. Chalupper and H. Fastl, "Dynamic Loudness Model (DLM) for normal and hearing-impaired listeners," Acta Acust. Acust. **88**, 378–386 (2002).

[19] P. Daniel and R. Weber, "Psychoacoustical roughness: Implementation of an optimized model," Acta Acust. Acust. **83**, 113–123 (1997).

[20] A. Camacho, "SWIPE: A sawtooth waveform inspired pitch estimator for speech and music," Ph.D. dissertation, University of Florida, Gainesville, FL, 2007.

[21] R. K. Mantiuk, A. Tomaszewska, and R. Mantiuk, "Comparison of four subjective methods for image quality assessment," Computer Graphics Forum **31**(8), 1–13 (1981).

[22] Ö. Axelsson, M. E. Nilsson, and B. Berglund, "A principal components model of soundscape perception," J. Acoust. Soc. Am. **128**(5), 2836–2846 (2010).

[23] J. Kang and M. Zhang, "Semantic differential analysis of the soundscape in urban open public spaces," Build. Environ. **45**(1), 150–157 (2010).

[24] J. Y. Jeon, P. J. Lee, J. Y. Hong, and D. Cabrera, "Non-auditory factors affecting urban soundscape evaluation," J. Acoust. Soc. Am. **130**(6), 3761–3770 (2011).

[25] R. Cain, P. Jennings, and J. Poxon, "The development and application of the emotional dimensions of a soundscape," Appl. Acoust. **74**(2), 232–239 (2013).

[26] A. Field, *Discovering Statistics using SPSS*, 2nd ed. (Sage, London, 2005), 816 pp.

[27] S. Siegel and N. J. Castellan, *Nonparametric Statistics for the Behavioral Sciences*, 2nd ed. (McGraw-Hill, New York, 1988), pp. 262–284.

[28] D. G. Kleinbaum and M. Klein, *Logistic Regression: A Self-learning Text*, 3rd ed. (Springer, New York, 2010), pp. 103–164.